Defence Research and Development Canada
Recherche et développement pour la défense Canada

DEFENCE R&D DÉFENSE

Spot SAR ATR Using Wavelet Features and Neural Network Classifier

N.M. Sandirasegaram

Defence R&D Canada – Ottawa
TECHNICAL MEMORANDUM
DRDC Ottawa TM 2005-154
October 2005

Canada

Spot SAR ATR Using Wavelet Features and Neural Network Classifier

N.M. Sandirasegaram
DRDC Ottawa

Defence R&D Canada – Ottawa

Technical Memorandum

DRDC Ottawa TM 2005-154

October 2005

106-04005/

© Her Majesty the Queen as represented by the Minister of National Defence, 2005

© Sa majesté la reine, représentée par le ministre de la Défense nationale, 2005

Abstract

An overview and performance summary of an Automated Target Recognition (ATR) algorithm based on spot Synthetic Aperture Radar (SAR) imagery is described in this report. Feature extraction and classification are very important steps in the ATR process. In this algorithm, the two dimensional wavelet decomposition method was applied to SAR targets to extract features. Selection of an appropriate mother wavelet was done by testing various wavelets and selecting the one which produced the smallest variation between features for the same target types, and the largest variation between features for different target types. After extensive testing, the Reverse Biorthogonal was selected as the best mother wavelet for this application. Second level approximation coefficients were used as features, and were fed into a Multi Layer Perceptron (MLP) neural network (NN) for classification. The MLP NN was trained using a supervised method, the standard delta rule. The classification results are shown using Receiver Operation Characteristic (ROC) curves and Confusion Matrices. The analysed result shows that the Reverse biorthogonal wavelet features are as good as two-dimensional Fast Fourier Transform features in the MSTAR (Moving and Stationary Target Acquisition and Recognition) dataset application. Results also show that including confusers (objects that the ATR algorithm is not intended to classify) in the training dataset reduces false alarm because the classifier has learned to reject confusers during the training process.

Résumé

Ce rapport présente un aperçu et un résumé des performances d'un algorithme de reconnaissance automatique des cibles (ATR) fondé sur l'imagerie RSO (radar à synthèse d'ouverture) ponctuelle. L'extraction et la classification des caractéristiques sont des étapes très importantes du processus ATR. Dans cet algorithme, la méthode de décomposition en ondelettes bidimensionnelles a été appliquée à des cibles RSO pour l'extraction de caractéristiques. On a sélectionné une ondelette mère appropriée en testant diverses ondelettes puis en sélectionnant celle qui produisait la plus petite variation entre les caractéristiques pour les mêmes types de cibles, et la plus grande variation entre les caractéristiques pour des types de cibles différents : après des essais à grande échelle, on a retenu l'ondelette inverse biorthogonale aux fins de l'application. Des coefficients d'approximation de deuxième niveau ont été utilisés comme caractéristiques et introduits dans un réseau neuronal perceptron multicouches (MLP) aux fins de la classification. Le réseau MLP a été entraîné à l'aide d'une méthode supervisée, soit la règle delta standard. Les résultats de la classification sont figurés au moyen de courbes FER (fonction d'efficacité du récepteur) et de matrices de confusion. Les résultats analysés montrent que les caractéristiques de l'ondelette inverse biorthogonale sont aussi bonnes que les caractéristiques TFR (transformée de Fourier rapide) 2D de l'application d'ensemble de données MSTAR (acquisition et reconnaissance de cibles mobiles et fixes). Ils montrent en outre que l'inclusion d'éléments de confusion (objets que l'algorithme ATR n'est pas censé classifier) dans l'ensemble de données d'entraînement réduit les fausses alarmes, car le classificateur apprend ainsi à rejeter les éléments de confusion lors du processus d'entraînement.

This page intentionally left blank.

Executive summary

Synthetic Aperture Radar (SAR) is a very useful imaging sensor for defense application because it operates at all times of day and under all the weather conditions. Because the image produced is based on the backscatter of a high frequency incident beam, however, targets located in imaging are often difficult to identify. Collection of radar imagery increases every year as new systems are being deployed with radar imaging capability (CP-140, Predator, etc.) and there is no time to go through the collected data manually. The intention of the Automated Target Recognition (ATR) method is to bring the attention of the image analyst to any potential targets that may exist in an image and to pass on all the known information about them to the image analyst. The known information can be target type, size, etc. The ATR process can be broken into many parts, of which feature extraction and classification are two of the most fundamental. Of the many feature extraction and classification algorithms that are available, the two dimensional wavelet feature extraction and Multi Layer Perceptron (MLP) Neural Network (NN) classification algorithms have been used successfully in many applications. These techniques can be extended for use with SAR imagery.

For the algorithm described in this document, the two dimensional wavelet algorithm is applied to SAR images to extract significant signatures (features) from each target. The MLP NN is then used to identify the types of a given target, based on these signatures. The optimal wavelet for feature extraction was determined by comparing the feature variability for similar and different target types produced by various wavelets. The Reverse Biorthogonal Wavelet was selected as the best since it produced the lowest feature variability for similar target types, and the highest variability for different target types. The ATR algorithm was implemented in Matlab using the two dimensional wavelet transform defined in the wavelet toolbox, and the MLP neural network was developed by DRDC Ottawa. To evaluate the ATR algorithm, Receiver Operation Characteristic (ROC) curves and confusion matrices were used. The ROC curve shows the relationship between the percentage of correct detections and percentage of false alarms. The confusion matrix shows the number of targets that were correctly recognized, misclassified and rejected.

The results show that including non-target vehicles ("confusers") in training set decreases the percentage of false alarm by more than 79%. This shows that inclusion of expected confusers in the training dataset will reduce the false alarm. The percentage of correct classification of detected targets was 86.4%. Therefore, wavelet feature extraction algorithm has some ability to separate different target types and it can be used in ATR system to improve recognition rate.

Future work includes the incorporation of these algorithms with other ATR algorithms to improve the classification performance. In addition to incorporating more than one feature extraction algorithms, the dominant features (best features that representing the target type) should be selected to reduce the dimensionality of the input to a classifier.

Sandirasegaram, N. (2005). Spot SAR ATR Using Wavelet Features and Neural Network Classifier. DRDC Ottawa TM 2005-154. Defence R&D Canada – Ottawa.

Sommaire

Le radar à synthèse d'ouverture (RSO) est un capteur d'imagerie très utile aux fins d'applications de défense, car il fonctionne à toute heure du jour et dans toutes les conditions météorologiques. Toutefois, l'image étant produite à partir de la rétrodiffusion d'un faisceau haute fréquence incident, les cibles imagées sont souvent difficiles à identifier. La collecte d'images radar s'accroît d'année en année à mesure que sont déployés de nouveaux systèmes dotés de l'imagerie radar (CP-140, Predator, etc.), et le temps manque pour l'examen manuel des données recueillies. La méthode ATR (reconnaissance automatique des cibles) vise à signaler à l'analyste toutes les cibles qui peuvent être indiquées sur une image et à lui transmettre toute l'information connue sur celles-ci : type, taille, etc. Le processus ATR peut se diviser en plusieurs parties, les deux plus importantes étant l'extraction et la classification. Parmi tous les algorithmes d'extraction et de classification existants, les algorithmes d'extraction de caractéristiques au moyen d'ondelettes bidimensionnelles et les algorithmes de classification au moyen de réseau neuronaux MLP se sont avérés efficaces dans de nombreuses applications. Ces techniques peuvent également être utilisées avec l'imagerie RSO.

En ce qui concerne l'algorithme décrit dans le présent document, l'algorithme utilisant l'ondelette bidimensionnelle est appliqué à des images RSO pour extraire des signatures d'intérêt (caractéristiques) de chaque cible. Le réseau neuronal MLP permet ensuite d'identifier les types d'une cible donnée en fonction de ces signatures. On a déterminé l'ondelette optimale pour l'extraction de signatures en comparant la variabilité des caractéristiques pour des types de cibles semblables et différentes produites par diverses ondelettes. L'ondelette inverse biorthogonale a été sélectionnée parce qu'elle produisait la plus faible variabilité des caractéristiques pour des types de cibles semblables, et la plus grande variabilité pour des types de cibles différents. L'algorithme ATR a été mis en oeuvre dans Matlab en utilisant la transformée d'ondelette bidimensionnelle définie dans la boîte à outils d'ondelettes, et le réseau neuronal MLP a été développé par RDDC Ottawa. Pour évaluer l'algorithme ATR, on a utilisé des courbes FER (fonction d'efficacité du récepteur) et des matrices de confusion. Les courbes ROC montrent la relation entre le pourcentage de détections correctes et le pourcentage de fausses alarmes. La matrice de confusion indique le nombre de cibles reconnues, mal classées et rejetées.

Les résultats indiquent que l'inclusion de véhicules non ciblés (éléments de confusion) dans l'ensemble de données d'entraînement réduit le pourcentage de fausses alarmes de plus de 79 %. Cela montre que le fait d'inclure des éléments de confusion prévus dans l'ensemble de données d'entraînement réduit le nombre des fausses alarmes. Le pourcentage de classification correcte des cibles détectées était de 86.4 %. Par conséquent, l'algorithme d'extraction utilisant les ondelettes permet dans une certaine mesure de séparer différents types de cibles et peut être mis à profit dans un système ATR pour améliorer le pourcentage de reconnaissance.

Les recherches futures comprennent l'intégration des algorithmes précités avec d'autres algorithmes ATR pour améliorer les performances de classification. Outre l'intégration de plusieurs algorithmes d'extraction de caractéristiques, il faudrait sélectionner les

caractéristiques dominantes (caractéristiques les plus représentatives du type de cible) de manière à réduire le volume des entrées dans un classificateur.

Sandirasegaram, N. (2005). Spot SAR ATR Using Wavelet Features and Neural Network Classifier. DRDC Ottawa TM 2005-154. R & D pour la défense Canada – Ottawa.

Table of contents

Abstract i

Résumé i

Executive summary iii

Sommaire iv

Table of contents vi

List of figures vii

List of tables viii

1.0 Introduction 1

2.0 Detected target 2

3.0 Preprocessing 2

4.0 Feature Extraction method 4

4.1 2D FFT feature extraction method 5

4.2 Two dimensional wavelet feature extraction method 5

4.2.1 Mother wavelet selection method 9

5.0 MLP Neural Network Classification 12

6.0 Results and Discussion 13

6.1 Experimented data set 13

6.2 Selection of Mother wavelet 14

6.3 Classification results and discussion 16

7.0 Conclusion 23

References 24

List of acronyms 27

List of figures

Figure 1: Block diagram of target recognition processing steps. 1

Figure 2: a) Example Real SAR Image Including Targets, and b) Detected target from (a)..... 2

Figure 3: Representation of slant range and ground range. 3

Figure 4: Standardization of target orientation to north facing position a) Target BTR-70 before correcting its orientation and b) after rotating the BTR-70 to north facing orientation. . 3

Figure 5: Size of reduced target chip is 64 by 64 pixels. 4

Figure 6: Selected Fourier Coefficients features' location for the T72. 5

Figure 7: Pyramidal tree structure two levels of 2D wavelet decomposition steps for images. . 8

Figure 8: Two level decomposition of discrete wavelet transform. a) Block representation of an image. b) Block Wavelet representation of second resolution levels. c) Original BMP2 image. d) The BMP Wavelet representation of second resolution levels. 9

Figure 9: Example – $\mu\sigma_2$ and $\mu\sigma_1$ plots of three different mother wavelets. (a) Three mother wavelet points (b) Distance between reference point and each mother wavelet point 11

Figure 10: Four layer Multi Layer Perceptron Neural Network. 12

Figure 11: $\mu\sigma_2$ and $\mu\sigma_1$ plots of tested mother wavelets. 15

Figure 12: ROC curve for the BMP2, BTR70 and T72 target types. 18

Figure 13: Optical and SAR images of Slicy target. (a) Optical image of Slicy target. (b) SAR image of Slicy target. 20

Figure 14: ROC curve for the BMP2, BTR70 and T72 target types – Slicy confusers are included in the training set. 20

Figure 15: Example of training cover area before and after inclusion of confusers in training set. (a) Before inclusion of confusers. (b) After inclusion of confusers. 21

Figure 16: ROC curve for the BMP2, BTR70 and T72 target types – 2S1 and D7 confusers are included in the training set. 22

List of tables

Table 1: Example - $\mu\sigma_2$ and $\mu\sigma_1$ values for three different mother wavelets. 11

Table 2: Mother wavelets tested for the selection of the best mother wavelet for classification. 12

Table 3: Details of the MSTAR targets used as the training Set. 14

Table 4: Details of MSTAR targets used as testing samples. 14

Table 5: Confusers used for ROC curve test. 14

Table 6: Result of feature extraction using various mother wavelets. 15

Table 7: Short distance from each mother wavelet ($\mu\sigma_1$, $\mu\sigma_2$) point to RP. 16

Table 8: Confusion Matrix for Pd of 1. 17

Table 9: Confusion Matrix at P_d is 0.9. 18

Table 10: MLP Confusion matrix for 2D FFT features at P_d is 0.9 ([25]). 19

Table 11: Confusion Matrix at lower false alarm and maximum detection rate. 19

Table 12: Confusion Matrix for Pd of 0.9 when Slicy confusers were included in the training set. 20

Table 13: Confusion Matrix at Pd of 0.9 – 2S1 and D7 confusers are included in the training set. 22

1.0 Introduction

A Synthetic Aperture Radar (SAR) system transmits microwaves and records the reflected signals from the scene being imaged. The SAR imaging technique is significantly different from optical techniques where the image acquired is based on the energy emitted by objects in the sensor's field of view (FOV). A complete description of SAR imaging can be found in [1,2,3]. The brightness of a pixel in a SAR image depends on strength of the reflected signal: high strength signals appear brighter than low strength signals. The strength of the reflected signal depends on many factors [4,5,6,7] such as size, topography, and water content of objects, as well as radar wavelength, signal polarization, and incident angle. The reflected signal's properties help Automatic Target Recognition (ATR) systems in identifying unknown targets. In recent years, the research community has increased the use of SAR imagery in ATR development [8,9,10]. SAR ATR systems, however, are still in the developmental stage and it will be some time until they are fully operational.

A SAR-based ATR system requires a fast and effective discriminator to detect a target and recognize the type of target from the radar return [11]. This study concentrates on the target recognition process. The target recognition process as illustrated in Figure 1, can be broken into meaningful processing steps: preprocessing, feature extraction and classification.

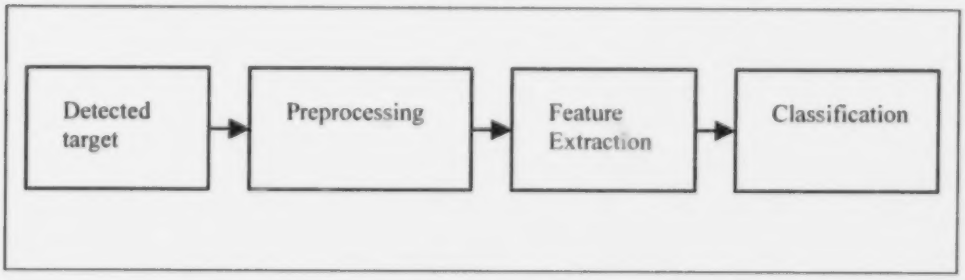


Figure 1: Block diagram of target recognition processing steps.

Pattern recognition systems have been developed and applied to many applications, including optical character recognition [12-16] and data mining [17-23]. Methods such as Neural Network (NN) [24-28] and wavelet transforms [29,30,31] are applied in SAR applications. Both techniques; NN and wavelet transforms were applied to SAR land area classification [32]. In this study, the two dimensional discrete wavelet transform (DWT) algorithm is used to extract features from SAR imagery and the Multi-Layer Percetpron (MLP) NN algorithm is used to classify targets based on the extracted features. Seven different mother wavelets were applied to extract wavelet features and then they were compared with each other to select the best mother wavelet for this application. Then the selected mother wavelet was used to extract features from SAR targets and MLP NN was used to classify using these features. Preprocessing, feature extraction, and classification will be discussed in the next three sections. Included in the discussion of feature extraction and classification is a discussion of the DWT feature extraction and MLP classification methods.

2.0 Detected target

A SAR scene may contain many man-made objects, as shown in Figure 2a. These objects can be extracted from the scene either by applying automated detection algorithm(s) or by manual detection. The selected object is called the detected target (target chip) because this object is the focus of attention of the image analyst and/or the detection algorithm(s) for further consideration. The target chip is shown in Figure 2b.

a) b)

Figure 2: a) Example Real SAR Image Including Targets, and b) Detected target from (a).

An ATR system receives the detected target and then processes it to determine its identity. The output of the system is a list of known target identifications and the confidence associated with these identifications. The testing and training target chips used in this investigation are 128 pixels by 128 pixels. These target chips are high quality SAR imagery of military ground vehicles, and are known as the MSTAR (Moving and Stationary Target Acquisition and Recognition) public data set.

3.0 Preprocessing

Some feature extraction algorithms and classification algorithms are sensitive to location shift, rotation, and intensity. Reducing the sensitivity to these geometric and radiometric variations can enhance the accuracy of an ATR system. The following is a description of the pre-processing techniques used in this study. The same methods were used by English [24].

For this study, only the value of magnitude was considered and the phase information ignored. All the target chips (training and testing) were converted from slant-range to ground-range using the following:

$$G = \frac{S}{\mathrm{Cos}(\theta)}, \qquad (1)$$

where S is slant range, G is the ground range, and θ is the depression angle. Figure 3 shows the relationship between these three parameters.

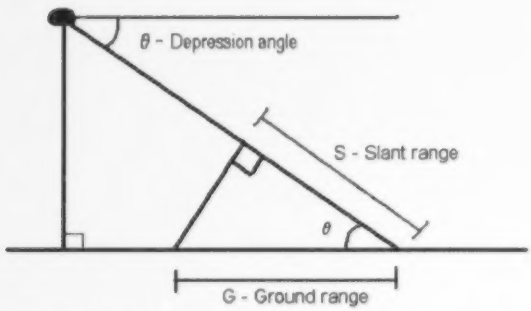
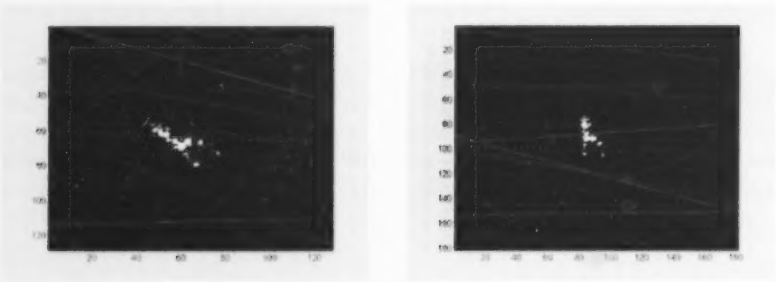


Figure 3: Representation of slant range and ground range.

In the MSTAR dataset the target contained within each chip is oriented independently of the other chips. To bring the targets into a standardized target orientation, each target was rotated to a vertical orientation with the front end of the target facing north. For this rotation, the orientation angle is found using ground truthing information. Figure 4a and Figure 4b show the target position before and after rotation, respectively.



a) b)

Figure 4: Standardization of target orientation to north facing position a) Target BTR-70 before correcting its orientation and b) after rotating the BTR-70 to north facing orientation.

After rotating to the standardized orientation, the highest energy reflecting point of the target was found in the target chip. A median filter was applied to isolate the target region in the target chip, and a search was subsequently used to locate the highest energy return point in that region. The energy at this point represents the highest energy reflected towards the radar sensor for particular sensor depression angle and target orientation. This point of highest energy was then used as the centre point for a new chip, the size of which 64 pixels by 64 pixels. The reduced size of the new target chip was more than enough to cover any target in the MSTAR 3 target problem. Figure 5 shows the new target chip cropped using the target chip shown in Figure 4b.

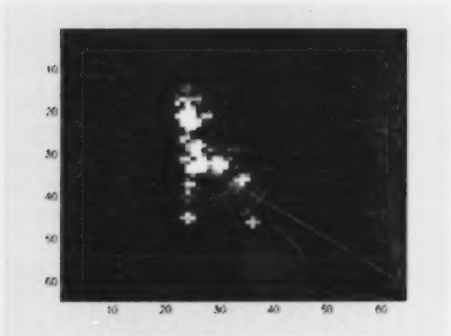


Figure 5: Size of reduced target chip is 64 by 64 pixels.

The final preprocessing step was to normalize the target chips. Normalization alters the pixel values such that the mean intensity is zero and the standard deviation value is one for each chip. This was done by subtracting the unsigned value of the mean intensity from each pixel intensity and dividing the result by the standard deviation of the target chip. The mathematical explanation is given in the following equation;

$$X'(i,j) = \frac{X(i,j) - \overline{X}}{\sigma_X}, \quad 1 \le i \le N, 1 \le j \le M, \qquad (2)$$

where X is the non-normalized target chip, X' is the normalized target chip, $\overline{X}$ is the mean intensity, σ_x is standard deviation of the target chip X, N is the number of pixels in the range direction, and M is the number of pixels in the cross range direction. The normalized target chip was then passed on to feature extraction algorithms.

4.0 Feature Extraction method

Feature extraction is one of the important steps in the ATR process. Feature extraction algorithms extract unique information or a signature from each target. A very good feature extraction algorithm gives smaller variation between the same type of targets and larger variation between different types of targets. Selection of a good feature extraction algorithm is important; otherwise, it will be difficult to differentiate between targets of different types and misclassification will occur. The two-dimenstional (2D) Fast Fourier Transform (FFT) feature extraction method, which was used in [25], is briefly described in Section 4.1. The Wavelet Transform (WT) feature extraction method was used for this study and a detailed description is given in Section 4.2.

4.1 2D FFT feature extraction method

In [25], the 2D FFT was used to extract the features of Fourier coefficients (FCs). In a 64 pixel by 64 pixel image, there are 4096 FCs, half of which are redundant. Therefore, only 2048 FCs actually represent the 64 pixel by 64 pixel image. Using all of these FCs features would slow down the classification and training processes. Therefore, the number of FCs was reduced to 16 for each target type, with different harmonics used for each target type. Therefore, the best 16 harmonics for each target type were combined. The selection procedure can be found in the [25]. In English's work [24], 256 FCs were selected for each target type, but the 256 harmonics of each target type were not combined. For example, the selected 256 FCs for T72 target type are displayed in Figure 6.

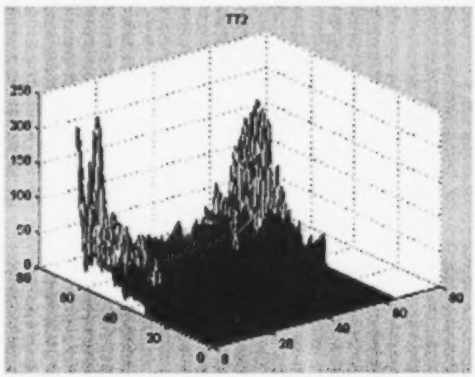


Figure 6: Selected Fourier Coefficients features' location for the T72.

The 2D Fourier analysis gives the frequency response of an image and depends on the periodic components that occur in the image. Fourier analysis offers good frequency resolution but not good space localization. Therefore, if good space resolution is required another method must be used. One solution is to use all pixels in the target chip as inputs to the classifier. However for a classifier such as the MLP NN, 4096 inputs (64 x 64) is too computationally intensive. In addition, this approach also provides the classifiers with both target and non-target information. This is problematic because non-target information will cause a false alarm or misclassification problem. An alternative approach is the WT as it provides both spatial and frequency localization [33,34,35].

4.2 Two dimensional wavelet feature extraction method

The WT decomposes the original image into several sub images of coarser resolution than the original image. At each level of decomposition, four sub images (LL, LH, HL, HH) are obtained according to the mother wavelet function. LL contains low frequency components in horizontal and vertical directions. LH contains low frequency components in horizontal direction and high frequency components in vertical direction. HL contains high frequency

components in horizontal direction and low frequency components in vertical direction. HH contains high frequency components in both horizontal and vertical directions. The resolution of all sub images is reduce to a size that is one-fourth the size of the original image. Fourier analysis expresses the original image in terms of a sum of bases functions. These bases functions are sinusoids of different frequencies. Similarly, the wavelet analysis expresses the original image in terms of sum of bases functions, but these base functions are shifted and scaled versions of mother wavelet. Here, we consider only the separable 2D discrete wavelet transform because it can be computed using 1D scaling and wavelet functions.

The 2D DWT decomposes an image in terms of wavelet and scaling functions [36]. The following equation shows the 2D wavelet decomposition of an image $X(u_1,u_2)$, $X(u) \in L^2(\Re)$.

$$X(u) = \sum_{k\in Z^2} a_{j_0,k}\phi^{LL}_{j_0,k}(u) + \sum_{b\in B}\sum_{j\geq j_0}\sum_{k\in Z^2} d^b_{j,k}\psi^b_{j,k}(u), \tag{3}$$

where $b \in B := \{LH, HL, HH\}$, $\phi^{LL}_{j_0,k}(u)$ is the 2D dilated and translated scaling function, a_{j_0} are the scaling or approximation coefficients, $\psi^b_{j,k}$ is the 2D translated and dilated wavelet function, $d^b_{j,k}$ are the detail or wavelet coefficients, j ($j \geq j_0$) is a scale factor, k (two dimensional variable $k=(k_1,k_2)$) is the shifting factor of the wavelet and scaling functions, respectively, and j_0 is a fixed scale. These 2D functions can be broken down into the product of 1D functions [36, 37]:

$$\phi^{LL}_{j_0,k}(u) = \phi_{j_0,k_1}(u_1)\,\phi_{j_0,k_2}(u_2), \tag{4}$$

$$\psi^{LH}_{j,k}(u) = \psi_{j,k_1}(u_1)\,\phi_{j_0,k_2}(u_2), \tag{5}$$

$$\psi^{HL}_{j,k}(u) = \phi_{j_0,k_1}(u_1)\,\psi_{j,k_2}(u_2), \tag{6}$$

$$\psi^{HH}_{j,k}(u) = \psi_{j,k_1}(u_1)\,\psi_{j,k_2}(u_2), \tag{7}$$

where $\phi_{j_0,k_1}(u_1)$ and $\psi_{j,k_1}(u_1)$ are the 1D column direction scaling and wavelet functions, and $\phi_{j_0,k_2}(u_2)$ and $\psi_{j,k_2}(u_2)$ are the 1D row direction scaling and wavelet functions. These 1D translated and dilated functions form the mother wavelet and scaling functions:

$$\phi_{j_0,n}(t) = 2^{j_0/2}\phi(2^{j_0}t - n), \tag{8}$$

$$\psi_{j,n}(t) = 2^{j/2}\psi(2^j t - n),\ n = k_1,k_2 \text{ and } t = u_1, u_2, \tag{9}$$

where ϕ is the mother scaling function and the ψ is the mother wavelet function. The approximation coefficients $a_{j_0,k}$ and the detail coefficients $d^b_{j,k}$ can be found using the following equations [38] in discrete form:

$$a_{j_0,(k_1,k_2)} = \sum_{u_1}\sum_{u_2} X(u_1,u_2)\,\phi_{j_0,k_1}(u_1)\,\phi_{j_0,k_2}(u_2) \Rightarrow \text{LL}, \tag{10}$$

$$d^{LH}_{j,(k_1,k_2)} = \sum_{u_1}\sum_{u_2} X(u_1,u_2)\,\psi_{j,k_1}(u_1)\,\phi_{j_0,k_2}(u_2) \Rightarrow \text{LH}, \tag{11}$$

$$d^{HL}_{j,(k1,k2)} = \sum_{u1}\sum_{u2} X(u_1,u_2)\,\phi_{j0,k1}(u_1)\,\psi_{j,k2}(u_2) \Rightarrow \text{HL}, \tag{12}$$

$$d^{HH}_{j,(k1,k2)} = \sum_{u1}\sum_{u2} X(u_1,u_2)\,\psi_{j,k1}(u_1)\,\psi_{j,k2}(u_2) \Rightarrow \text{HH}. \tag{13}$$

The 2D approximation and detail coefficients can be calculated from the 1D scaling and wavelet functions. This separable 2D discrete wavelet transform can be computed by applying 1D low-pass and the high-pass digital filters [36,37,38,39] to an image $X(u_1,u_2)$. The low-pass filter represents the scaling function and high-pass filter represents the wavelet function [34]. This decomposition using 1D filters is achieved by first filtering the rows of the input image with the low pass and high pass filters and decimating the result by 2. The columns of the two decimated images are then filtered with the low and high pass filters and are decimated by 2 again. This first-level decomposition produces the four filtered images discussed above: LL, LH, HL, and HH.

The processes required for two levels of wavelet decomposition are shown in Figure 7. LL and LLLL are level 1 and level 2 approximation images. LH, HL and HH are the level 1 detail images, and LLLH, LLHL and LLHH are the level 2 detail images. The level 2 approximation image contains low frequency information in both the vertical and horizontal directions, that is high frequency noise is removed in LLLL. In this study, the level 2 approximation sub-band image (LLLL) is fed into the classifier for identification. The Matlab wavelet toolbox is used to extract the second level approximation image. The sub-band image block diagram and an example illustration the 2D WT decomposition are shown in Figure 8.

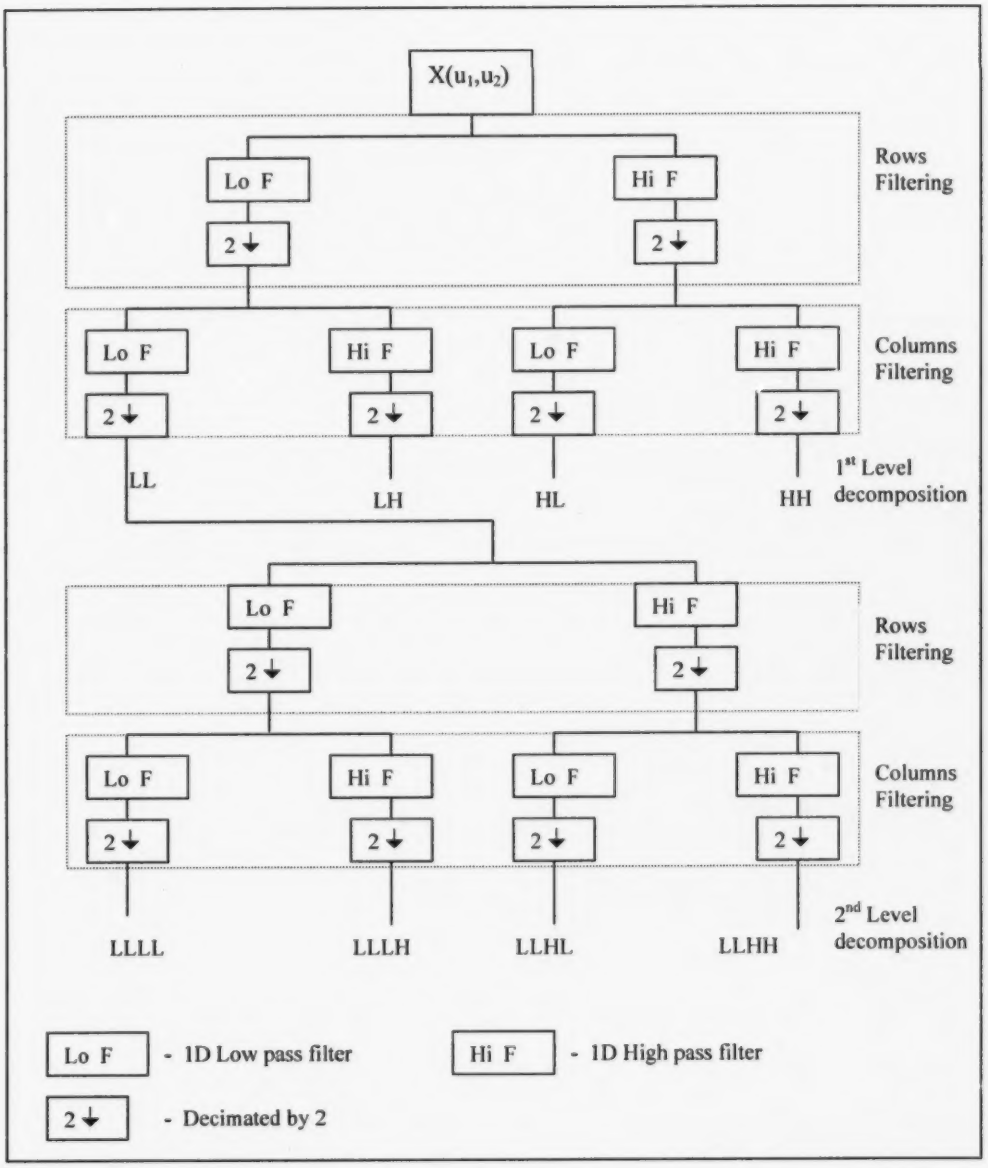


Figure 7: Pyramidal tree structure two levels of 2D wavelet decomposition steps for images.

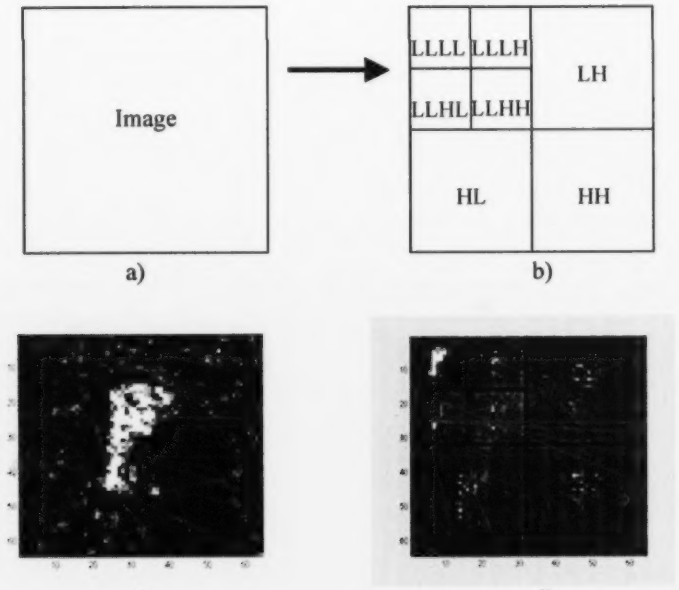


Figure 8: Two level decomposition of discrete wavelet transform. a) Block representation of an image. b) Block Wavelet representation of second resolution levels. c) Original BMP2 image. d) The BMP Wavelet representation of second resolution levels.

In Fourier transformation, images breakdown to sinusoidal and cosine functions of various frequencies but in the wavelet transformation images breakdown to translated and scaled versions of the mother wavelet. Many different numbers of wavelet families exist and the only difference between these mother wavelets are on shape and duration of the waveform. Within each family of wavelets, the wavelets are classified by the number of vanishing moments (order number). This indicates the smoothness of the wavelet and flatness of the frequency response of the wavelet filters. But all the mother wavelets do not have vanishing moments. We have investigated seven mother wavelets and five of them had vanishing moments.

4.2.1 Mother wavelet selection method

A second level wavelet decomposition was applied to the training set using the Biorthogonal spline, Coiflet, Daubechies, discrete approximation of Meyer, Haar, Reverse biorthogonal, and Symlet mother wavelets one by one to obtain the approximation coefficients. The standard deviation method was then applied to the approximation coefficients obtained for each wavelet in order to determine the best mother wavelet. Variations of features in the

same type of target ($\mu\sigma_1$) and variations of mean features between different types of target ($\mu\sigma_2$) were measured for the selection process. A small value of $\mu\sigma_1$ indicates the features in the same type of target are invariant, and a large value of $\mu\sigma_2$ indicates the feature distances between different target types are far a part from each other. The mother wavelet which maximizes the distance ($\mu\sigma_2$ - $\mu\sigma_1$) is deemed the best of wavelet.

The mathematical expressions involved with computation of $\mu\sigma_1$ are as follows:

$$\mu\sigma_1 = \frac{1}{N * M}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}, \tag{14}$$

$$\sigma_{ij} = \sqrt{\frac{1}{L}\sum_{s=1}^{L}(x_s - \mu_{ij})^2}, \tag{15}$$

$$\mu_{ij} = \frac{1}{L}\sum_{s=1}^{L}x_s \tag{16}$$

In (14) N is the number of target types (3 in this case), M is the number of features (256 in this case), and σ_{ij} is the standard deviation of each feature for each target type. In (15) L is the number samples in feature j and target type i, x_s is the wavelet coefficients of the s^{th} sample, j^{th} feature and target type i, and μ_{ij} is the mean value of target type i and feature j. To find out the variation of features in different target types, the standard deviation ($\sigma\mu_j$) was calculated using each feature mean value as shown equation 17. Then mean of the standard deviation ($\mu\sigma_2$) was calculated to determine the average variation of features in different target types using equation 19.

$$\sigma\mu_j = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mu_{ij} - \overline{\mu_j})^2}, \tag{17}$$

$$\overline{\mu_j} = \frac{1}{N}\sum_{i=1}^{N}\mu_{ij}, \tag{18}$$

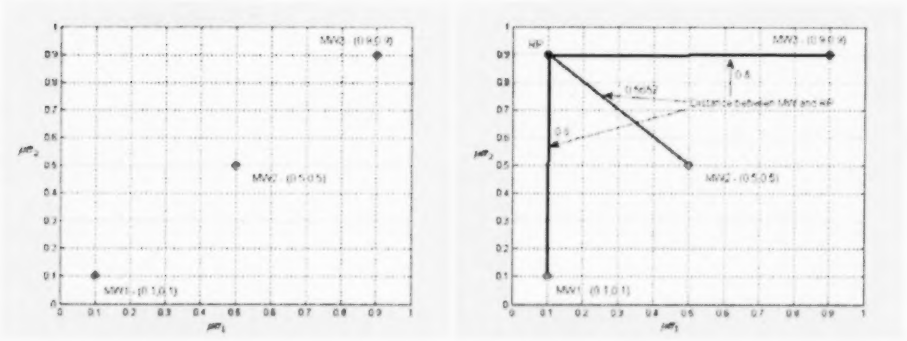
$$\mu\sigma_2 = \frac{1}{M}\sum_{j=1}^{M}\sigma\mu_j \tag{19}$$

In (19) $\overline{\mu_j}$ is the mean value of this j^{th} feature of all the target types, and the remaining variables are as defined above. The best mother wavelet was selected as the one that produced the largest variance between different targets types with the same feature and the smallest variance between the same target types with the same feature. For example, assume $\mu\sigma_2$ and $\mu\sigma_1$ have been calculated for three different mother wavelets, as shown in Table 1. Figure 9(a) shows three points plotted as $\mu\sigma_2$ versus $\mu\sigma_1$ for the three wavelets (MW1, MW2 and MW3). At the point MW1, values of $\mu\sigma_2$ and $\mu\sigma_1$ are small, indicating that this is not a good feature because it is very difficult to distinguish between two different target types. The

point MW3 has larger values for both $\mu\sigma_2$ and $\mu\sigma_1$, this will cause a misclassification problem. The point MW2 is a compromise on the inter-target class variability ($\mu\sigma_2$) and the intra-target class variability ($\mu\sigma_1$). Therefore, MW2 is the best of the three mother wavelets. A minimum distance method can be used to determine the best mother wavelet. After selecting a reference point (RP) such as minimum point of $\mu\sigma_1$ and maximum point of $\mu\sigma_2$ from the experimented mother wavelets, the distance between each point ($\mu\sigma_1$, $\mu\sigma_2$) and the RP is computed. The mother wavelet with the minimum distance is considered the best. In the case of MW1, MW2 and MW3 the minimum distance occurred for MW2, thus supporting the original observation that MW2 was the best wavelet. Figure 9(b) shows the distances computed for each mother wavelet. This distance method was followed for selecting the best wavelet in this study.

Table 1: Example - $\mu\sigma_2$ and $\mu\sigma_1$ values for three different mother wavelets.

MOTHER WAVELET	$\mu\sigma_2$	$\mu\sigma_1$
MW1	0.1	0.1
MW2	0.5	0.5
MW3	0.9	0.9



(a) (b)

Figure 9: Example – $\mu\sigma_2$ and $\mu\sigma_1$ plots of three different mother wavelets. (a) Three mother wavelet points (b) Distance between reference point and each mother wavelet point

The mother wavelets, which were subject to this experiment, are listed in Table 2. Features were extracted using each mother wavelet in Table 2 in order to calculate $\mu\sigma_1$ and $\mu\sigma_2$.

Table 2: Mother wavelets tested for the selection of the best mother wavelet for classification.

	MOTHER WAVELETS	ORDER
1	Biorthogonal spline	Order 1.1, 3.1 and 6.8 are used in this analysis
2	Coiflets	Order 1, 3 and 5 are used in this analysis
3	Daubechies	Order 2, 24 and 45 are used in this analysis
4	Discrete approximation of Meyer	Order 3 is used in this analysis
5	Haar	Order 1 is used in this analysis
6	Reverse biorthogonal	Order 1.3, 3.1 and 6.8 are used in this analysis
7	Symlets	Order 1, 10 and 20 are used in this analysis

5.0 MLP Neural Network Classification

Neural Networks are characterized by two phases; the learning phase and the responding phase. The learning process can be achieved using either supervised or unsupervised learning methods. In supervised learning methods, the inputs and desired outputs are known to the NN. The learning process adjusts weights (parameters) to get desired outputs from inputs. In unsupervised learning methods, only the inputs are known to the NN. The learning process adjusts weights to group similar patterns of inputs to same output nodes. The number of groups depends on number nodes that exist in the output layer. In this work, the supervised method is followed. During the learning phase, the NN learns the relationship between the information given at the input and the information requested from the output. In the responding phase, test data similar to the learning input data is fed to the input. Based on the learned response the NN then predicts the response to the test data.

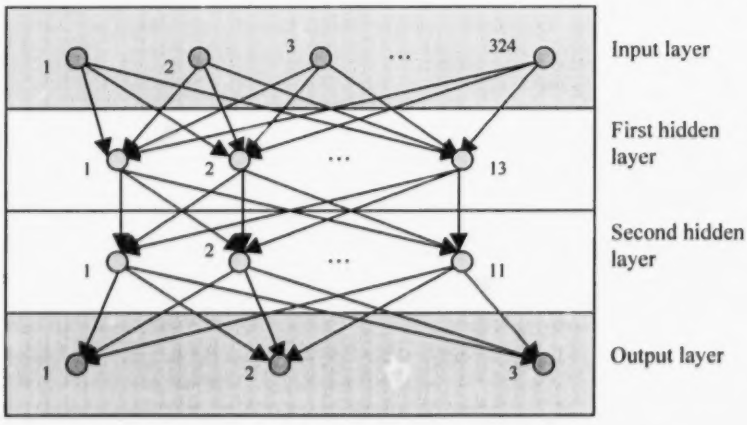


Figure 10: Four layer Multi Layer Perceptron Neural Network.

The most commonly used nonlinear regression model is the MLP NN [42], which is capable of learning nonlinear function mappings [41]. The MLP NN has been used more extensively in various problems [43] than any other NN [40]. The typical MLP NN is composed of an input layer, an output layer, and at least one hidden layer [40]. In this work, the MLP neural network used a standard back-propagation algorithm called, the delta rule algorithm [44], to update the weights. Details of the derivation for the training and testing algorithms are not discussed here. The MLP NN was configured as illustrated in Figure 10, with 13 nodes in the first hidden layer, 11 nodes in the second hidden layer, 256 nodes in the input layer and 3 nodes in the output layer. The number of input nodes depends on the number of features chosen. In this work, 256 features were chosen; therefore, 256 nodes should be included in the input layer of the NN. Because in this application the NN was use to classify the three military ground vehicles, a BMP-2, BTR-70 and T-72, three nodes were required in the output layer. Node 1 was assigned to T-72, Node 2 was assigned to BTR-70, and Node 3 was assigned to BMP-2. The number of hidden layers and the number of nodes in each hidden layer were decided using empirical testing of the NN; these are not optimized numbers.

6.0 Results and Discussion

As mentioned in section 2.0, various mother wavelets were applied to the public MSTAR data set to determine the best mother wavelet for this application. 256 features were then extracted from each target chip using the selected mother wavelet. Features extracted from the training data were used to train the MLP NN, and features extracted from the testing data were used for testing the trained MLP NN. This section starts with brief discussion about the experiment data, followed by the selection of mother wavelets, and the classification results of the 2D wavelet features.

6.1 Experimented data set

The United States (US) Defense Advanced Research Projects Agency (DARPA) has made part of the MSTAR data set available to the public. The MSTAR public data set contains many spotlight SAR vehicle images including 10 types of former Soviet Union vehicles, with the azimuthal angle ranging between 0 and 360 degrees, and depression angles of 15 and 17 degrees. Three types of vehicles (BMP-2, BTR-70 and T-72) are investigated in this analysis. All three types of vehicles, which were imaged at a 17-degree depression angle, and all azimuthal angles, were used for training the MLP NN. The target types with serial number, and the number of samples used for training, are listed in the Table 3. Similarly, all three types of vehicles, which were imaged in 15 degree and all azimuthal angles, were used for testing the 2D wavelet feature extraction and the MLP NN classifier. The target types with serial numbers and the number of samples used for testing are listed in the Table 4.

Table 3: Details of the MSTAR targets used as the training Set.

TARGETS TYPE AND SERIAL NUMBER	# OF SAMPLES	COMMENTS
T-72 (132)	232	All the targets collected at 17 degree depression angle, full aspect coverage and 30 cm resolution
BTR – 70 (c72)	233	
BMP – 2 (9563)	233	
Total	= 698	

Table 4: Details of MSTAR targets used as testing samples.

TARGETS TYPE AND SERIAL NUMBER	# OF SAMPLES	COMMENTS
T-72 (812)	195	All the targets collected at 15 degree depression angle, full aspect coverage, and 30 cm resolution
T-72 (s7)	191	
T-72 (132)	196	
BTR-70 (c72)	196	
BMP-2 (9563)	195	
BMP-2 (9566)	196	
BMP-2 (c21)	196	
Total	= 1365	

Confuser targets were used to generate Receiver Operation Characteristic (ROC) curves for evaluate the classifier performance. Two types of confusers were used, the 2S1 self-propelled howitzer, and the D7 bulldozer. The confuser targets information is listed on the Table 5.

Table 5: Confusers used for ROC curve test.

TARGETS TYPE AND SERIAL NUMBER	# OF SAMPLES	COMMENTS
2S1	274	All the confusers collected at 15 degree depression angle, full aspect coverage, and one foot resolution
D7	274	
Total	= 548	

6.2 Selection of Mother wavelet

The Matlab wavelet toolbox was used to extract the wavelet coefficients from each image. Different mother wavelets were tested and the rows of $\mu\sigma_1$ and $\mu\sigma_2$ were calculated for each wavelet as per equations 14 and 19. The calculated values are listed in Table 6. These points are plotted in Figure 11 as per the method described in Section 4.2.1. The RP was determined using the minimum and maximum values of $\mu\sigma_1$ and $\mu\sigma_2$ respectively. From the Table 6, the minimum value of $\mu\sigma_1$ is 0.999 and maximum value of $\mu\sigma_2$ is 1.3471, therefore, the RP is (0.999, 1.347) and it is denoted as RP on Figure 11. The distance from each point in Figure 11 to the RP was calculated and is listed in Table 7. An examination of the distances shows that the minimum distance corresponds to the Reverse Biorthogonal Wavelet Pairs with order of 3.1. Thus, this wavelet was selected for further analysis.

Table 6: Result of feature extraction using various mother wavelets.

MOTHER WAVELETS	# OF SAMPLES		$\mu\sigma_2$	$\mu\sigma_1$
Name	***Id***	***Order***		
Biorthogonal spline	P1	1.1	0.391	1.514
	P2	3.1	0.587	3.458
	P3	6.8	0.245	1.097
Coiflets	P4	1	0.337	1.375
	P5	3	0.241	1.063
	P6	5	0.222	1.021
Daubechies	P7	2	0.352	1.428
	P8	24	0.538	2.020
	P9	45	**1.347**	5.887
Discrete approximation of Meyer	P10	N/A	0.363	1.519
Haar	P11	N/A	0.391	1.514
Reverse Biorthogonal Wavelet Pairs	P12	1.3	0.315	1.319
	P13	3.1	0.293	1.130
	P14	6.8	0.233	1.014
Symlets	P15	1	0.391	1.514
	P16	10	0.226	1.014
	P17	20	0.214	**0.999**

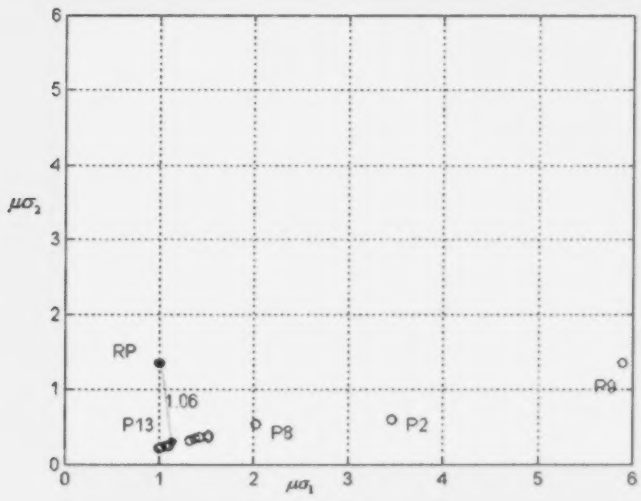


Figure 11: $\mu\sigma_2$ and $\mu\sigma_1$ plots of tested mother wavelets.

Table 7: Short distance from each mother wavelet ($\mu\sigma_1$, $\mu\sigma_2$) point to RP.

MOTHER WAVELETS		DISTANCE FROM RP
Name	***Order***	
Biorthogonal spline	1.1	1.086
	3.1	2.573
	6.8	1.107
Coiflets	1	1.077
	3	1.108
	5	1.125
Daubechies	2	1.084
	24	1.302
	45	4.888
Discrete approximation of Meyer	N/A	1.113
Haar	N/A	1.086
Reverse Biorthogonal Wavelet Pairs	1.3	1.081
	3.1	**1.062**
	6.8	1.114
Symlets	1	1.086
	10	1.121
	20	1.133

6.3 Classification results and discussion

Section 6.2 showed that, of the wavelets tested, the Reverse Biorthogonal Wavelet had the best performance. Therefore, this wavelet was used to extract features from the target images. The two-hidden-layer MLP NN classifier was then used to classify the targets. First the wavelet feature extraction algorithm was applied to the training data set (Table 3) and then the MLP NN was trained on the extracted features so it could learn to distinguish between target types. After training, the performance of the MLP NN was evaluated using the testing set (Table 4) and confuser set (Table 5). The evaluation was conducted using the confusion matrix and the ROC curves. A complete explanation of this evaluation technique can be found in [45], which is a NATO working paper. In the ROC curve, the y-axis represents the number of targets correctly detected divided by the total number of same type of targets presented to the classifier (as a percentage), and the x-axis represents the number of confusers detected divided by the total number of confusers presented to the classifier (as a percentage). These points are calculated by changing the output threshold value of the MLP NN classifier from 0 to 1 in increments of 0.01.

There are three output nodes in the MLP NN's output layer, and each node represents a particular target type (Node 1 for T-72, Node 2 for BTR-70 and Node 3 for BMP-2). Assigned to each node is a particular threshold value. When a target is tested, each of the

three output nodes of the NN produces a value between 0 and 1. The output value of each node is then compared to the threshold value of that node. A target is considered similar to a trained target type if the output value is greater than or equal to the threshold value, and dissimilar to a trained target type if the output value is less than the threshold value. In the event that two output nodes accept a target, the node with the greatest difference between the calculated and defined threshold values wins.

One of the experiments was conducted by setting the threshold value for each output to zero. This ensures the classifier will classify the target into one of the class types, because output values will always be between 0 and 1 inclusive. This decision indicates which tested targets are more similar to which of the trained target types. Table 8 shows the confusion matrix. The asterisks in the table indicate vehicle types with serial number, used for training. The percentage of correct classification rate (P_{CC}) is higher for the vehicle types (same serial number) used in training compared to the other targets. For each target, the correct recognition rate is shown in bold. Overall, the P_{CC} is 84.2%. According to the result in the table, the confusers D7 and 2S1 are more similar to the T72 vehicle type.

Table 8: Confusion Matrix for Pd of 1.

TARGETS TYPE AND SERIAL NUMBER	BMP2	BTR70	T72	P_{CC} (%)
BMP2 (SN_9563)	**144**	7	44	73.85
BMP2 (SN_9566)	**146**	6	44	74.49
BMP2 (SN_C21)*	**185**	0	11	94.39
BTR70 (SN_C71)*	7	**185**	4	94.39
T72 (SN_132)*	7	1	**188**	95.92
T72 (SN_S7)	36	8	**147**	76.96
T72 (SN_812)	39	2	**154**	78.97
2S1	69	60	145	-
D7	21	1	252	-
Overall Pcc				84.18

The percentage of false alarm rate (P_f) was calculated using the confuser targets in Table 5 and the percentage of correct detection rate (P_d) was calculated using the testing targets in Table 4. The P_f and P_d rates were computed for different output threshold values and then plotted as P_d vs. P_f as shown in Figure 12. In this plot, the lower left and the upper right corner points were generated at the output threshold value of 1 and 0 respectively, and the rest of the points were generated at threshold values between 0 and 1. This graph is called a ROC curve. The P_f and the P_d were calculated separately for each target type. For a random classifier, the values of the P_d and P_f are approximately equal for different output threshold values (the area under the curve of a random classifier is approximately 0.5). The broken line shown in the Figure 12 is the random classifier line. Any curve below this line is not good because the P_d is less than the P_f. In this study all the classifiers produced ROC curves above

the random classifier line. The performance of the classifier is best when the tested targets belong to the same class of targets used in training.

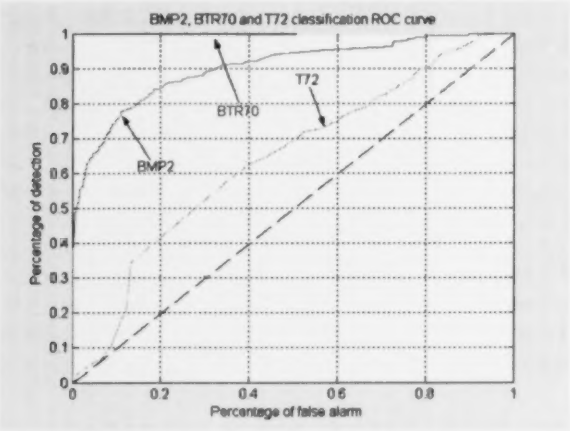


Figure 12: ROC curve for the BMP2, BTR70 and T72 target types.

The threshold values at the output nodes were adjusted to ensure a P_d of 0.9. Typically a P_d of 0.9 is used in the MSTAR evaluation method as a standard operating point [25,46]. Table 9 shows the confusion matrix that was produced for this P_d. The ratio of the total number of correct classifications to the total number of targets detected ($P_{cc/d}$) is highest when the serial number of the test target is the same as the serial number of the trained target (BMP2 (SN_C21), BTR70 (SN_C71), and T72 (SN_132)). To compare the wavelet features with 2D FFT features, classification based on the 2D FFT features was performed. The results for a P_d of 0.9 are shown in the Table 10. The overall $P_{cc/d}$ is slightly higher for the wavelet-based features. The false alarm rate is lower for the wavelet-based features (P_f is 0.85 at P_d of 0.9) than the 2D FFT-based features (P_f is 0.89 at P_d of 0.9).

Table 9: Confusion Matrix at P_d is 0.9.

Targets type and serial number	BMP2	BTR70	T72	Others
BMP2 (SN_9563)	**110**	4	35	46
BMP2 (SN_9566)	**122**	2	39	33
BMP2 (SN_C21)*	**178**	0	10	8
BTR70 (SN_C71)*	1	**178**	3	14
T72 (SN_132)*	5	0	**187**	4
T72 (SN_S7)	29	7	**140**	15
T72 (SN_812)	31	1	**147**	16
2S1	31	47	131	65
D7	10	1	244	19
				Overall $P_{cc/d}$ = 86.41%

Table 10: MLP Confusion matrix for 2D FFT features at P_d is 0.9 ([25]).

Targets type and serial number	BMP2	BTR70	T72	Others
T-72(812)	**141**	13	15	26
T-72(S7)	**122**	13	30	26
T-72(132)	**177**	3	5	11
BTR-70(C72)	3	**183**	0	10
BMP-2(9563)	5	5	**170**	15
BMP-2(9566)	31	16	**121**	28
BMP-2(C21)	23	4	**152**	17
2S1	45	93	87	49
D7	22	2	238	12
				Overall $P_{cc/d}$ = 85.20%

An ideal classifier's ROC curve should go through the point (0,1), which represents a zero false alarm rate and a 100% detection rate. Therefore, the point, which is closest to (0,1), is the best threshold value available to get a higher detection rate and a lower false alarm rate. The best threshold values found were 0.9906 for node 1 (T-72), 0.9937 for node 2 (BTR-70), and 0.9805 (BMP-2). The confusion matrix produced based on these values and the results are listed in the Table 11. The P_d is 62.3% and the P_f is 41.2%. The overall Pcc/d is 93.2%.

Table 11: Confusion Matrix at lower false alarm and maximum detection rate.

Targets type and serial number	BMP2	BTR70	T72	Others
BMP2 (SN_9563)	76	0	9	110
BMP2 (SN_9566)	84	0	10	102
BMP2 (SN_C21)*	151	0	5	40
BTR70 (SN_C71)*	1	134	2	59
T72 (SN_132)*	1	0	167	28
T72 (SN_S7)	15	2	90	84
T72 (SN_812)	13	0	90	92
2S1	5	15	74	180
D7	3	1	128	142
				Overall Pcc/d = 93.18%

Although the P_f has been reduced, it is still quite high. To further reduce the P_f rate, slicy confusers (shown in the Figure 13) were included in the training set and the MLP NN was retrained. The slicy confuser is a man-made object and this ATR system was not intended to detect or classify this slicy confuser. The trained MLP NN was evaluated using the testing set data from Table 4 and confusers from Table 5. A ROC curve and the confusion matrix for a P_d of 0.9 were generated. The ROC curve is shown in the Figure 14 and the confusion matrix is shown in the Table 12.

(a) (b)

Figure 13: Optical and SAR images of Slicy target. (a) Optical image of Slicy target. (b) SAR image of Slicy target.

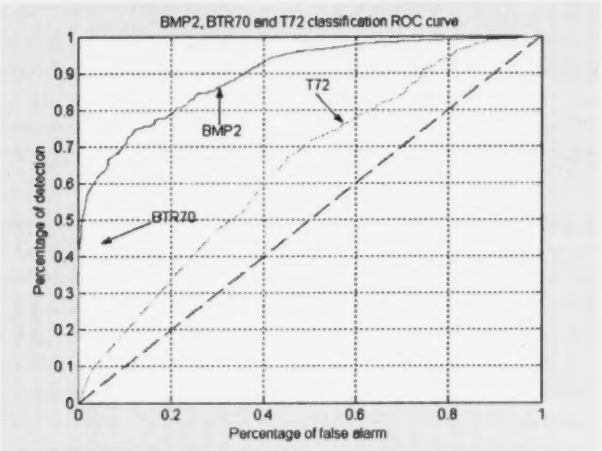


Figure 14: ROC curve for the BMP2, BTR70 and T72 target types – Slicy confusers are included in the training set.

Table 12: Confusion Matrix for Pd of 0.9 when Slicy confusers were included in the training set.

Targets type and serial number	BMP2	BTR70	T72	Others
BMP2 (SN_9563)	**124**	5	35	31
BMP2 (SN_9566)	**122**	5	50	19
BMP2 (SN_C21)*	**172**	0	13	11
BTR70 (SN_C71)*	1	**181**	3	11
T72 (SN_132)*	3	0	**186**	7
T72 (SN_S7)	27	6	**134**	24
T72 (SN_812)	23	0	**139**	33
2S1	42	68	116	48
D7	15	1	219	39
				Over all Pcc/d = 86.09%

The ROC curve shown in Figure 14 shows no improvement over that in Figure 12, and the overall $P_{cc/d}$ from Table 12 (86.1%) is actually less than that shown in Table 10 (86.46%). The reason for this result is explained as follows. Assume 'x' represents BMP-2, 'o' represents T-72 and '@' represents BTR-70, as shown in Figure 15. Before including confusers in the training set, the trained MLP NN covers a small area for each type of target (Area 1 for target type BMP-2, Area 2 for target type T-72, and Area 3 for BTR-70) as shown in Figure. After the inclusion of confusers in the training set, the trained MLP NN covers a different area, as shown in Figure 15b. Consider a target tested under both cases. The tested target belongs to BMP-2 type and it is marked in blue color in the Figure 15. The tested target is located inside the Area 1 in Figure 15a and the MLP NN classifier will recognize it correctly. In the Figure 15b, the tested target is located out side the Area 1 and the classifier will recognize it as an other or unknown target type. Consequently, the overall $P_{cc/d}$ will be reduced when confursers are included in the training set.

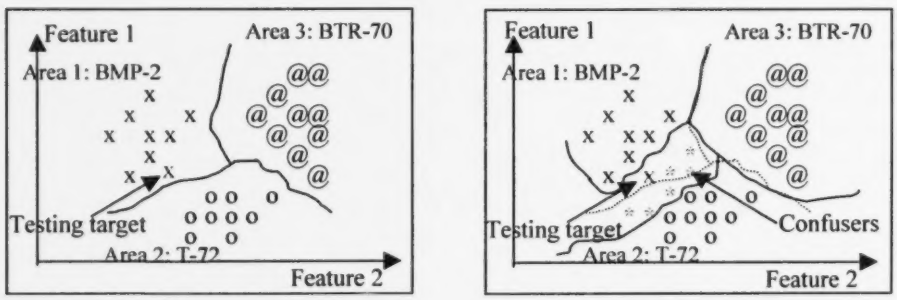


Figure 15: Example of training cover area before and after inclusion of confusers in training set. (a) Before inclusion of confusers. (b) After inclusion of confusers.

Including different type of confusers (Slicy) in the training set reduces the false alarm rate slightly. Therefore, the same types of confusers, imaged at a 17^0 depression angle, were used in the training set to train the MLP NN to further reduce the false alarm rate. The test set confusers was imaged at a 15^0 depression angle. Again the MLP NN was trained and then evaluated using the ROC curve and confusion matrix at 0.9 P_d. The ROC curve is shown in Figure 14 and the confusion matrix is shown in Table 13. This ROC curve is clearly better than the previous ROC curves. That is, these curves pass closer to the ideal ROC curve than the previous ROC curves. The confusion matrix shows that the overall Pcc/d is 84.7%, which is less than that of the previous $P_{cc/d}$ values. However, the P_f value is reduced to 5.7%. This shows that if similar type of confusers are added to the training set then the false alarm rate will be reduced.

Table 13: Confusion Matrix at Pd of 0.9 – 2S1 and D7 confusers are included in the training set.

Targets type and serial number	BMP2	BTR70	T72	Others
BMP2 (SN_9563)	**124**	9	47	15
BMP2 (SN_9566)	**116**	7	45	28
BMP2 (SN_C21)*	**186**	0	8	2
BTR70 (SN_C71)*	1	**188**	4	3
T72 (SN_132)*	4	0	**192**	0
T72 (SN_S7)	19	4	**121**	47
T72 (SN_812)	36	4	**114**	41
2S1	3	7	14	250
D7	3	1	3	267
				Overall Pcc/d = 84.70%

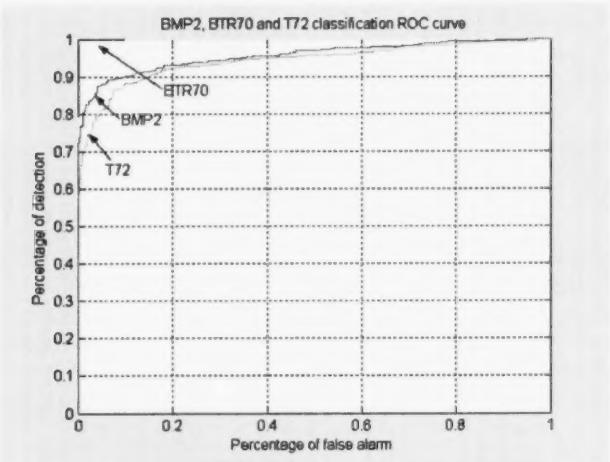


Figure 16: ROC curve for the BMP2, BTR70 and T72 target types – 2S1 and D7 confusers are included in the training set.

7.0 Conclusion

In this study comparison of seven different mother wavelets were conducted using the 3-class MSTAR SAR dataset. The Reverse Biorthogonal Wavelet Pairs mother wavelet performed better than all the other wavelets tested based on the following combinations; minimum variation within the same class type and maximum variation between different class types. Compared to the 2D FFT features, second level wavelet decomposition features only showed marginal improvement in classification performance. The MLP NN classifier has better detection and lower false alarm than the random classifier while using the Reverse Biorthogonal Wavelet Pairs features.

The MLP NN classifier was trained with three different combinations of datasets. One dataset contained all the training samples (Table 3), second dataset contained all the training samples and slicy confuser, and third dataset contained all the training samples and similar confusers as tested confusers but imaged in different depression angle. The correct classification rate has not varied noticeably for all the three datasets but the false alarm ratio was highly reduced in the third dataset and slightly reduced in the second dataset compared to first dataset. This experiment shows that adding the expected confusers in the training set will reduce the false alarm ratio.

Future work in this area may be the fusion of features produced by FFT and WT feature extractors. By fusion of these algorithms, many numbers of features will be extracted from each target. This will give high variations between the targets of the same class and low variations between the targets of the different classes. Therefore, we have to use few numbers of features but dominant features from the both feature extraction methods. In this study, only second level of approximation coefficients were considered and the details of coefficients were ignored. Therefore, WT features should be investigated in more details to extract meaningful features for SAR ATR application.

References

1. Henderson F. M., Lewis A. J., and Ryerson R. A. (1998). Manual of Remote Sensing: Principles and Applications of Imaging Radar. 3rd ed. Volume 2. John Wiley & Sons, Inc.
2. Oliver C., and Quegan S. (1998). Understanding Synthetic Aperture Radar Images. Artech House.
3. Franceschettti G., and Lanari R. (1999). Synthetic Aperture Radar Processing. CRC Press.
4. Freeman T. (January 26, 1996). What is Imaging Radar?. (online) NASA/Jet Propulsion Laboratory. http://southport.jpl.nasa.gov/desc/imagingradarv3.html (April 6, 2004).
5. Mouginis-Mark P. Introduction to Radar Remote Sensing (Page 2): Radar Backscatter as a Function of Incidence Angle. (online) Mouginis-Mark P. http://satftp.soest.hawaii.edu/space/hawaii/vfts/kilauea/radar_ex/page2.html (6 April 2004).
6. (1996) Remote Sensing. (online) Japan Association of Remote Sensing All rights reserved. http://www.profc.udec.cl/~gabriel/tutoriales/rsnote/contents.htm (6 April 2004).
7. InfoPACK User Guide Version 1.0. (online) InfoSAR Ltd. http://www.infosar.co.uk/docs/infopack.html (April 7, 2004).
8. D A Giglio, ed., Algorithms for Synthetic Aperture Radar Imagery I & II, Proceedings of SPIE, 1994 & 1995.
9. E G Zelnio and R J Douglass, ed., Algorithms for Synthetic Aperture Radar Imagery III, Proceedings of SPIE, 1996.
10. E G Zelnio, ed., Algorithms for Synthetic Aperture Radar Imagery IV - IX, Proceedings of SPIE, 1997 - 2003.
11. Chen-Pang,Y., Choongyeun, C., and Jeffrey, S.H. (2003). Toward target recognition from synthetic aperture radar imagery using electromagnetics-based signatures, *The International Society for Optical Engineering*, Vol. 42, 2129-2149.
12. Wilmott C. (online) Optical Character Recognition Using Distance Measurements on Polygonal Approximations. http://www.cs.byuh.edu/research/wilmott/491/paper.html (April 7, 2004).
13. Belaïd A. (online) OCR: Print. http://cslu.cse.ogi.edu/HLTsurvey/ch2node5.html#SECTION232 (April 7, 2004).
14. Xu, Y., and Nagy, G. (1999). Prototype Extraction and Adaptive OCR. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(12), 1280- 1296.
15. Mello C.A.B., and Lins R.D. (1999). A Comparative Study on Commercial OCR Tools. *Vision Interface'99*, 224- 323. Quebéc.
16. Pujari, A.K., Naidu, C.D., Jinaga, B.C. (2002). An Adaptive Character Recognizer for Telugu Scripts Using Multiresolution Analysis and Associative memory, *Third Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 2002)* [electronic journal], URL: http://www.ee.iitb.ac.in/~icvgip/PAPERS/319.pdf.
17. Vezzosi, S., Bedini, L. and Tonazzini, A., An integrated system for the analysis and the recognition of characters in ancient documents. (online). http://dienst.isti.cnr.it/Dienst/Repository/2.0/Body/ercim.cnr.iei/2002-TR-04/pdf?tiposearch=cnr&langver= (25 June 2004).

18. Kamath Chandrika, (December 8, 2003). Research Topics in Large Scale Data Mining. (online) Lawrence Livermore National Laboratory. http://www.llnl.gov/casc/sapphire/research.html (17 June, 2004).
19. Aas K., Huseby R. B. and Thune M.(1999). Data Mining: A Survey. (NR Report No. 942). Research and development at Norsk Regnesentral.
20. Tsantis, L. and Castellani, J. (Fall 2001). Enhancing Learning Environments Through Solution-based Knowledge Discovery Tools: Forecasting for Self-perpetuating Systemic Reform. (online) Journal of special education technology, http://jset.unlv.edu/16.4/tsantis/first.html (22 June, 2004).
21. Shanmugasundaram, J., Nagendra Prasad, M. V., Vadhavkar, S. and Gupta, A. (1999). Use of Recurrent Neural Networks for Strategic Data Mining of Sales Information. *1999 Information Resources Management Association International Conference (IRMA '99)*, 16-19 May 1999, Hershey, PA.
22. Li, T., Li, Q., Zhu, S. and Ogihara, M. (2002). A Survey on Wavelet Applications in Data Mining. *IGKDD Explorations Newsletter*, 4(2), 49-68.
23. Fabian Morchen, F. (2003). Time series feature extraction for data mining using DWT and DFT. http://www.informatik.uni-marburg.de/~databionics/papers/Moerchen2003b.pdf (25 June, 2004).
24. English, R.A. (2001). Automatic Target Recognition Using HNeT. (DREO TM 2001-080). Defence R&D Canada – Ottawa, Ottawa.
25. Sandirasegaram N.M. (2002). Automatic Target Recognition in SAR Imagery using a MLP Neural Network. (DRDC Ottawa TM 2002-120). Defence R&D Canada – Ottawa.
26. Zhao, Q., Xu, D. and Principe, J.C. (1998). , Pose Estimation For Sar Automatic Target Recognition. *Proceedings of the DARPA Image Understanding Workshop (IUW'98)*, 827-831. Monterey, California.
27. Gagnon, L. and Klepko, R. (1998). Hierarchical classifier design for airborne SAR images of ships, *Proceedings of SPIE Automatic Target Recognition V*, Vol. 3371, Orlando: SPIE Aerosense 98.
28. Kanellopoulos, I. A Survey of the Application of Neural Networks in Remote Sensing. (online) European Commission. http://europa.eu.int/en/comm/eurostat/research/supcom.95/16/result/node9.html#SECTION00070000000000000000 (23 June, 2004).
29. Kil, D. (1999). Technology Overview. (online) Rockwell Science Center http://www.killmine.com/OnlineManual/TechnologyOverview.htm (25 June 2004).
30. De Bonet, J. S., Viola, P., and Fisher III, J. W. (1998). Flexible Histograms: A Multiresolution Target Discrimination Model. *Proceedings of SPIE*, Volume 3370. Orlando.
31. Zhao, Q., Principe, J., Brennan, V., Xu, D., and Wang Z. (2000), Synthetic aperture radar automatic target recognition with three strategies of learning and representation. *The International Society for Optical Engineering*, 39(5), 1230-1244.
32. Qin, Q., Gillies, R. R., Lu, R., and Chen, S. (2004) An Integration of Wavelet Analysis and Neural Networks in Synthetic Aperture Radar Image Classification, *XXth ISPRS Congress*, p. 181 ff, Istanbul, Turkey.
33. Prasad, L., and Iyengar, S. (1997). Wavelet Analysis with applications to image processing. New York: CRC Press, p 279.
34. Zhang, B., Zhang, H., and Ge, S. (2004). Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memories, *IEEE Transactions on Neural Networks*, 15(1), 166-177.

35. Zhu, Z., Machiraju, R., Fry, B., and Moorhead, R. (1997). Wavelet-based Multiresolutional Representation of Computational Field Simulation Datasets, *Visualization '97*, 151-158. Arizona: IEEE.
36. Romberg, J., Choi, H., and Baraniuk, R. (2001). Bayesian Tree-Structured Image Modeling Using Wavelet-Domain Hidden Markov Models. IEEE Transactions on image processing, 10(7), 1056-1068.
37. Kopp, M (1995). Lossless Wavelet Based Image Compression with Adaptive 2D Decomposition. (TR-186-2-95-11). Institute of Computer Graphics and Algorithms, Vienna University of Technology, Austria.
38. Kotteri, K. A. Optimal, Multiplierless Implementations of the Discrete Wavelet Transform for Image Compression Applications, Master's Thesis, Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, April 2004.
39. Mallat, S. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Transactions on pattern analysis and machine intelligence, 11(7), 674-693
40. Anderson, D., McNeill, G. (1992). Artificial Neural Networks Technology. (Rome Laboratory RL/C3C, Griffiss AFB, NY, F30602-89-C-0082). Kaman Sciences Corporation.
41. Nikolaev, N. CIS 311 Neural Networks. (Online) Goldsmiths College, University of London. http://homepages.gold.ac.uk/nikolaev/cis311.htm (22 Sep. 2004).
42. Sarle, W. S. (1994), Neural Networks and Statistical Models, Proceedings of the Nineteenth Annual SAS Users Group International Conference, 1538-1550,NC, USA: SAS Institute Inc.
43. Franco, L., and Cannas, S.A.(200), Generalization and Selection of Examples in Feed Forward Neural Networks, Neural Computation, 12 (10), 2420-2449.
44. Lippmann, R. P. (1987), An introduction to computing with neural networks, IEEE ASSP Mag., 4(2), 4-22.
45. English, R.A. (2003). Classifier Evaluation Methodology using the MSTAR Public Data Set. DRDC Ottawa SL 2003-014. Defence R&D Canada - Ottawa.
46. Ross, T., Worrell, S., Velten, V., Mossing, J., and Bryant, M. (1998). Standard SAR ATR Evaluation Experiments Using the MSTAR Public Release Data Set. Proc. SPIE Conf.

List of acronyms

1D	One Dimensional
2D	Two Dimensional
ATR	Automatic Target Recognition
DARPA	Defense Advanced Research Projects Agency
DRDC	Defence Research and Development Canada
DWT	Discrete Wavelet Transform
FC	Fourier coefficient
FFT	Fast Fourier Transform
FOV	Field Of View
NATO	North Atlantic Treaty Organization
NN	Neural Network
MLP	Multi Layer Perceptron
MSTAR	Moving and Stationary Target Acquisition and Recognition
ROC	Receiver Operation Characteristic
RP	Reference Point
SAR	Synthetic Aperture Radar
WT	Wavelet Transform

This page intentionally left blank.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

Defence R&D Canada – Ottawa
3701 Carling Avenue
Ottawa, ON, K1A 0Z4

2. SECURITY CLASSIFICATION (overall security classification of the document, including special warning terms if applicable)

UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C or U) in parentheses after the title.)

Spot SAR ATR Using Wavelet Features and Neural Network Classifier (U).

4. AUTHORS (Last name, first name, middle initial)

Sandirasegaram, Nicholas, M

5. DATE OF PUBLICATION (month and year of publication of document)

October 2005

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)

35

6b. NO. OF REFS (total cited in document)

46

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

TECHNICAL MEMORANDUM

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

DRDC OTTAWA
3701 Carling Avenue
Ottawa, ON, K1A 0Z4

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)

13dt04

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

DRDC Ottawa TM 2005-154

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

An overview and performance summary of an Automated Target Recognition (ATR) algorithm based on spot Synthetic Aperture Radar (SAR) imagery is described in this report. Feature extraction and classification are very important steps in the ATR process. In this algorithm, the two dimensional wavelet decomposition method was applied to SAR targets to extract features. Selection of an appropriate mother wavelet was done by testing various wavelets and selecting the one which produced the smallest variation between features for the same target types, and the largest variation between features for different target types. After extensive testing, the Reverse Biorthogonal was selected as the best mother wavelet for this application. Second level approximation coefficients were used as features, and were fed into a Multi Layer Perceptron (MLP) neural network (NN) for classification. The MLP NN was trained using a supervised method, the standard delta rule. The classification results are shown using Receiver Operation Characteristic (ROC) curves and Confusion Matrices. The analysed result shows that the Reverse biorthogonal wavelet features are as good as two-dimensional Fast Fourier Transform features in the MSTAR (Moving and Stationary Target Acquisition and Recognition) dataset application. Results also show that including confusers (objects that the ATR algorithm is not intended to classify) in the training dataset reduces false alarm because the classifier has learned to reject confusers during the training process.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Automatic Target Recognition, ATR, Synthetic Aperture Radar imagery, spot SAR, MSTAR, Mulit-Layer Perceptron, MLP, Neural Network, Wavelet.

Defence R&D Canada

Canada's leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

www.drdc-rddc.gc.ca